

Parsing the Phonological Loop: Activation Timing in the Dorsal Speech Stream Determines Accuracy in Speech Reproduction

Alexander B. Herman,¹ John F. Houde,² Sophia Vinogradov,³ and Srikantan S. Nagarajan¹

¹Biomagnetic Imaging Laboratory, Department of Radiology and Biomedical Imaging, and Departments of ²Otolaryngology–Head and Neck Surgery and ³Psychiatry, University of California, San Francisco, San Francisco, California 94143

Despite significant research and important clinical correlates, direct neural evidence for a phonological loop linking speech perception, short-term memory and production remains elusive. To investigate these processes, we acquired whole-head magnetoencephalographic (MEG) recordings from human subjects performing a variable-length syllable sequence reproduction task. The MEG sensor data were source localized using a time–frequency optimized spatially adaptive filter, and we examined the time courses of cortical oscillatory power and the correlations of oscillatory power with behavior between onset of the audio stimulus and the overt speech response. We found dissociations between time courses of behaviorally relevant activations in a network of regions falling primarily within the dorsal speech stream. In particular, verbal working memory load modulated high gamma power in both Sylvian–parietal–temporal and Broca’s areas. The time courses of the correlations between high gamma power and subject performance clearly alternated between these two regions throughout the task. Our results provide the first evidence of a reverberating input–output buffer system in the dorsal stream underlying speech sensorimotor integration, consistent with recent phonological loop, competitive queuing, and speech–motor control models. These findings also shed new light on potential sources of speech dysfunction in aphasia and neuropsychiatric disorders, identifying anatomically and behaviorally dissociable activation time windows critical for successful speech reproduction.

Introduction

Verbal reproduction of heard speech sequences requires the coordination of the perceptual, short-term memory, and motor systems for speech. Dysfunction in verbal reproduction of speech sequences is observed in various aphasias, in stuttering, and in neuropsychiatric disorders such as schizophrenia (Barch, 2005; Baldo and Dronkers, 2006), but the neural substrates of these dysfunctions remain poorly understood. Although significant evidence supports our understanding of the components of speech reproduction individually, their dynamic integration at the level of large-scale neural circuits remains elusive. Speech perception, encompassing the spectrotemporal analysis and subsequent mapping of speech sounds onto stored sublexical (i.e., syllabic) representations, occurs in bilateral superior temporal gyrus

(STG) and superior temporal sulcus (STS). These sensory representations are mapped onto articulatory counterparts in Broca’s region/ventral premotor cortex (PMv), putatively via sensorimotor transformation in a functional area known as area Sylvian–parietal–temporal (Spt) in the left posterior planum temporale (PTp)/supramarginal gyrus (SMG) region, along a pathway known as the “dorsal stream” (Hickok and Poeppel, 2007; Saur et al., 2008). The dorsal stream also features prominently in current models of verbal working memory, which posit the existence of two buffers for the representation of speech sounds between perception and production—a perceptual buffer and a motor buffer—and identify an articulatory rehearsal process operating between the two for memory maintenance (Monsell, 1987; Vallar, 2006; Jacquemot et al., 2007; Baddeley, 2010). fMRI and lesion studies suggest an area located in the left inferior parietal lobe (IPL)/posterior STG as the most likely site for the perceptual phonological buffer, with Broca’s area/PMv the most likely location for the motor buffer for production (Bohland and Guenther, 2006; Jacquemot et al., 2007; Papoutsis et al., 2009; Rauschecker and Scott, 2009; Hickok et al., 2011). Once a verbal response is initiated, the brain maps the articulatory representation maintained in Broca’s area/PMv onto effectors via primary motor cortex.

Our understanding of the interactions among the seemingly separate neural processes that comprise heard speech reproduction has suffered from an inability to reconstruct cortical activity continuously through speech perception, maintenance, and re-

Received March 25, 2012; revised Dec. 7, 2012; accepted Dec. 18, 2012.

Author contributions: A.B.H., J.F.H., S.V., and S.S.N. designed research; A.B.H. performed research; A.B.H., J.F.H., and S.S.N. contributed unpublished reagents/analytic tools; A.B.H. analyzed data; A.B.H., J.F.H., S.V., and S.S.N. wrote the paper.

This work was funded in part by National Institutes of Health (NIH) Grants R01 DC4855, DC6435, DC10145, and NS67962, NIH/National Center for Research Resources University of California, San Francisco—Clinical and Translational Science Institute Grant UL1 RR024131, and National Science Foundation Grant BCS 0926196. We thank Anne Findlay, Susanne Honma, Tracy Luks, Corby Dale, Stephanie Aldebot, R. Alison Adcock, Leighton Hinkley, Adrian Guggisberg, Sarang Dalal, Johanna Zumer, and the NUTMEG development team.

Correspondence should be addressed to Srikantan S. Nagarajan, Biomagnetic Imaging Laboratory, 513 Parnassus Avenue, S362, Department of Radiology and Biomedical Imaging, University of California, San Francisco, San Francisco, CA 94143. E-mail: sri@ucsf.edu.

DOI:10.1523/JNEUROSCI.1472-12.2013

Copyright © 2013 the authors 0270-6474/13/335439-15\$15.00/0

sponse with high temporal and spatial resolution. In particular, the degree to which specific functions, such as input and output buffering or perceptual and motor image formation, dissociate between regions and across time remains unresolved. We utilized the millisecond time resolution and whole-brain cortical coverage of magnetoencephalography (MEG), combined with the subcentimeter spatial resolution offered by recent improvements in source localization algorithms, to examine the power fluctuations of neural oscillations during a variable-length syllable sequence reproduction task (Dalal et al., 2008). We hypothesized that syllable encoding and speech preparation would result in spectral changes consistent with known sensory and motor-related oscillatory processes, such as β power decreases coupled with more focal high gamma power ($H\gamma P$) increases, localized within the speech motor control network (Fukuda et al., 2010). Specifically, we predicted that $H\gamma P$ fluctuations in Spt and Broca's area/PMv would show the greatest syllabic load effect and would be correlated with behavioral performance in speech repetition, consistent with their roles as phonological and articulatory buffers within the speech motor control network.

Materials and Methods

Subjects. Seventeen right-handed healthy volunteers, 12 males and five females, participated in a verbal repetition task while undergoing magnetoencephalography recordings. Two of these subjects were subsequently excluded from our analysis because of inadequate performance. A high-resolution structural magnetic resonance image (MRI) was also obtained for each subject in a separate session. Subjects were screened for neurological conditions as well as contraindications for MEG and MRI. Written consent was obtained from each subject before the experiment. All experimental procedures were approved by the Institutional Review Board at the University of California, San Francisco.

Tasks. We designed a task similar to previous word-length effect tasks but with less of an explicit emphasis on rehearsal memory, in which subjects listened to and then repeated vocalized two- or four-syllable utterances after a brief delay. This task design minimized the confounding effects of lexical and syntactic processes and allowed us to elucidate the time course of neural circuit activation in natural speech repetition and the neural-behavioral correlates of syllable cognitive load as a test for phonological working memory-buffer functionality (Bohland and Guenther, 2006; Papoutsis et al., 2009; McGettigan et al., 2011). Subjects completed a verbal repetition task consisting of 80 two-syllable target trials and 80 four-syllable target trials pseudorandomly ordered, while lying supine undergoing MEG recording (Fig. 1A). Stimuli were prerecorded from a single female speaker and consisted of permutations of the syllables /ba/, /da/, and /pa/ (e.g., /ba da/ or /ba da pa ba/). On each trial of the experiment, subjects listened to either a two- or four-syllable target presentation, waited for a visual cue presented at a jittered delay uniformly distributed between 2050 and 2150 ms poststimulus onset, and then vocally repeated the stimulus pattern. Subjects had no previous knowledge or cues about the contents of the upcoming trial. Syllables lasted 470 ms on average and were separated within a trial by 50 ms. Within two-syllable trials, no syllable was repeated, and within four-syllable trials, no syllable pair was repeated. After the go cue, subjects had up to 3 s to complete their response, after which the experiment proceeded to the next trial. Correct and incorrect syllable repetition was manually recorded by the experimenters. Trials were labeled correct if subjects repeated the correct syllables in the correct order, within 3 s of the go cue.

Recordings. Tasks were administered while subjects underwent whole-head MEG recording in a 278-channel CTF Omega 2000 Biomagnetometer with third-order gradient correction (VSM MedTech), at a 1200 Hz sampling rate. Signals from radio-emitting coils, placed at the nasion and on both left and right sides of the head 1 cm rostral to the periauricular point, were triangulated to determine the position of the head relative to

the sensor array. These head-location points were then coregistered to high-resolution anatomical MRIs of the subjects' brains through a multiple-sphere head model. MRI scan sessions in which head movement exceeded 2 mm were discarded and repeated. Experiments and imaging were performed entirely at the University of California, San Francisco.

Analyses. Data were epoched into -1 to 7 s trials relative to onset of the first syllable. Channels and trials with high-frequency activity consist-

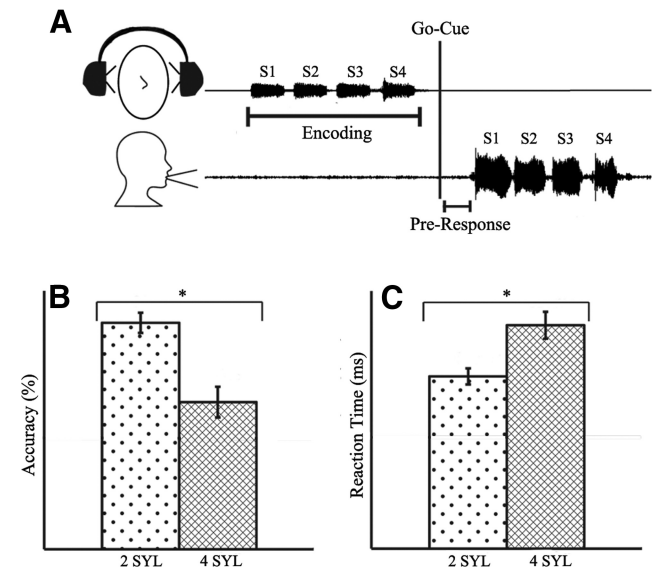


Figure 1. Experimental schematic and behavioral results. **A**, Digitized patterns of sample syllable stimulus and subject vocal response. **B**, Subjects responded correctly on an average of 73 ± 2 of 80 two-syllable trials (91%; 2 SYL) and 50 ± 6 of 80 four-syllable trials (63%; 4 SYL), and accuracy rates were significantly different between the two conditions ($p < 0.05$). **C**, Average response latencies for two- and four-syllable conditions differed significantly ($p < 0.05$), at 640 ± 50 and 830 ± 60 ms, respectively.

Table 1. Activations during syllable encoding: encoding phase

Time (ms)	Band (Hz)	Area	MNI	MNI coordinates			<i>t</i> value	<i>p</i> value
				<i>x</i>	<i>y</i>	<i>z</i>		
38	4–13	L STSp	BA22	–60	–38	8	5.0	0.001
38	50–120	L PMv	BA6	–53	–10	25	2.4	0.001
63	30–50	L STGm	BA22	–45	–18	0	–4.9	0.002
113	4–13	R TTG	BA41	65	–20	13	5.4	0.001
163	50–120	R SPL	BA7	10	–70	63	3.8	0.001
213	30–50	R PsCG	BA2	60	–25	38	–3	0.002
213	50–120	L Csm	BA6	–30	–5	50	3.6	0.001
263	13–30	L PMd	BA6	–52	13	52	–6.78	0.002
263	13–30	L STGa	BA21	–45	8	–8	–6.6	0.002
263	13–30	R ITL	BA20	68	–30	28	–3.8	0.002
263	13–30	R IFG	BA44	63	15	0	–3.8	0.002
338	30–50	R SMA	BA6	5	5	55	–6.8	0.001
338	50–120	R MCngt	BA31	18	–23	45	–5.1	0.002
463	50–120	R SMA	BA6	20	15	73	–4.2	0.002
488	30–50	L PMm	BA9/6	–45	5	43	–5.2	0.002
688	30–50	R SMA	BA6	–8	5	70	–6	0.001
838	13–30	R Csm	BA4	65	–28	43	–4.5	0.002
838	30–50	R PMd	BA4	50	–15	65	–5.1	0.002
863	4–13	R PTP	BA40	60	–65	28	4.5	0.001
863	13–30	L PMd	BA8	–33	18	53	–6.1	0.002
863	30–50	L PMm	BA6	–40	3	43	–6.6	0.002
963	50–120	L PMm	BA6	–43	–15	38	3.3	0.001

Time points, frequency bands, region label, Brodmann's areas, MNI coordinates, *t* values, and uncorrected *p* values are listed chronologically for peak voxels and time-window center points for statistically significant activations (FDR corrected, $p < 0.05$). Times are given relative to stimulus. R, Right hemisphere; L, left hemisphere.

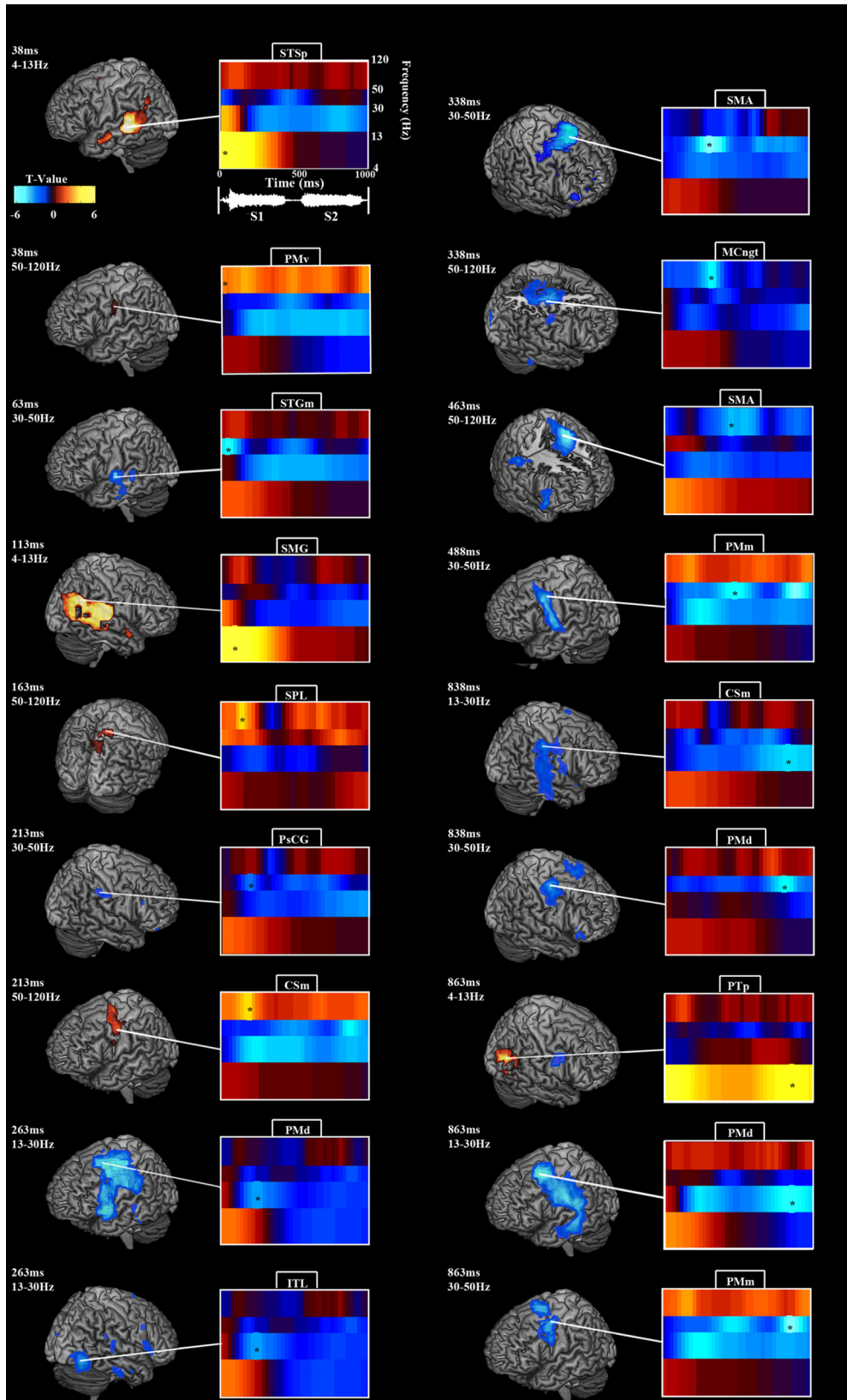


Figure 2. Oscillatory modulations during syllable encoding. All brain-rendering/spectrogram pairs follow the pattern detailed in the example in the top left corner. Each brain rendering depicts statistically significant activations for a single time point in a particular frequency band, indicated by an asterisk on the accompanying spectrogram. The spectrogram, in turn, displays the power time courses across frequency bands for the peak voxel in the brain rendering, indicated by a white line. Time 0 marks the onset of the first syllable, which ends at ~470 ms. The second syllable begins at ~520 ms and ends at ~990 ms. θ/α power peaked over bilateral auditory areas early, followed by β and $L\gamma P$ decreases over bilateral auditory and premotor areas and $H\gamma P$ increases over left premotor cortex. For details, see Results and Table 1.

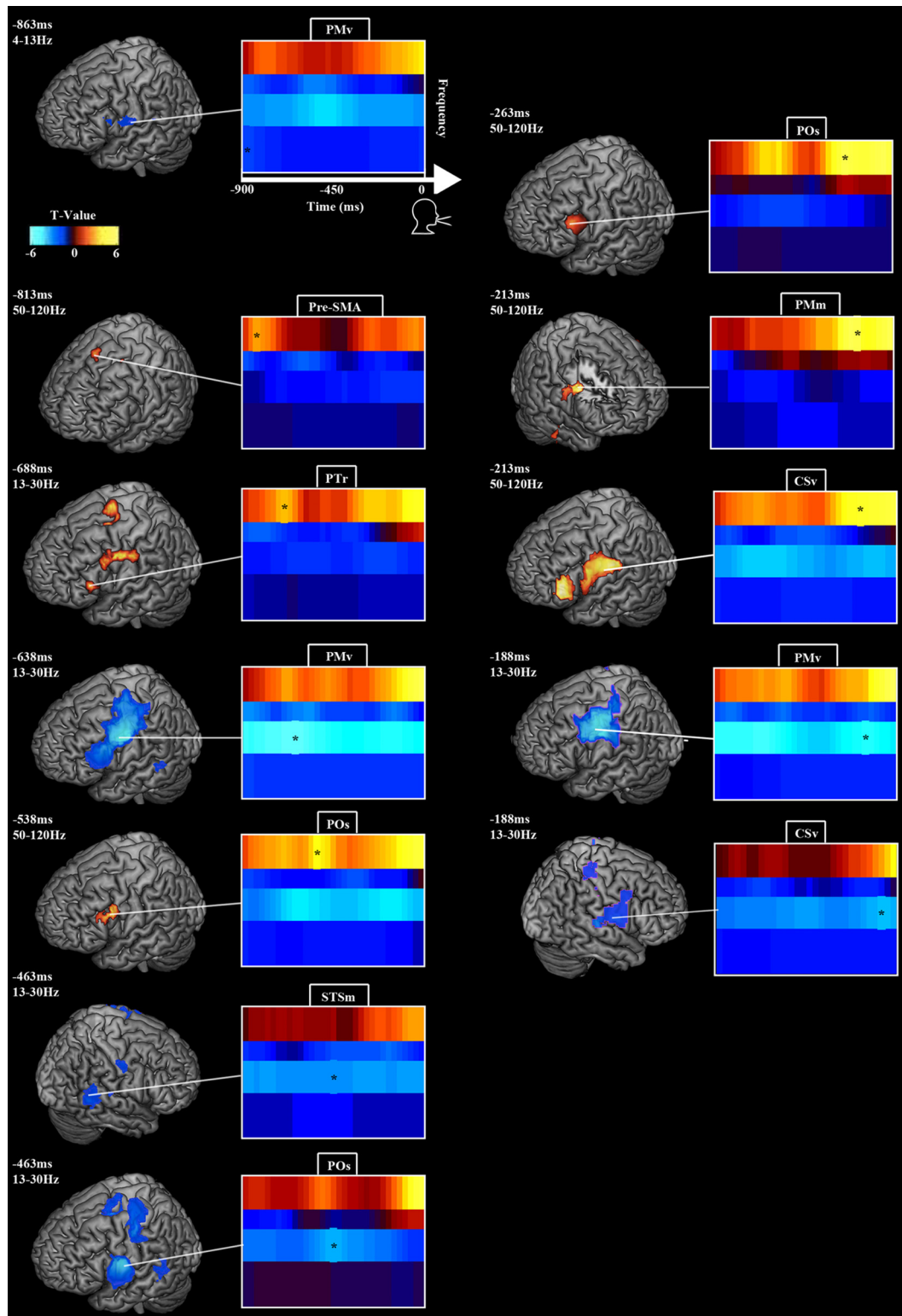


Figure 3. Oscillatory modulations during syllable speech preparation. Figure layout follows Figure 2. Time 0 marks voice onset. β power decreased and H γ P increased over PTR and SMA early in the pre-response period, followed by H γ P in POs, and as voice onset approached bilateral β power decreased and H γ P increased in PM. For details, see Results and Table 2.

tently >1.5 pT or in which the participant spoke during this interval were discarded. Only correct trials were analyzed. To enable neural-source localization, high-resolution anatomical MRIs were obtained for each subject and spatially normalized to a standard MNI template brain using SPM2 (<http://www.fil.ion.ucl.ac.uk/spm/software/spm2/>). Tomographic volumes of potential dipolar source locations (voxels) were generated from these normalized MRIs and coregistered to the MEG sensor arrays. To

avoid mislocalizations attributable to temporally correlated sources between the two hemispheres, data from sensors covering each hemisphere were analyzed both separately and together (Dalal et al., 2008, 2011). We did not subdivide sensor groups within hemisphere to avoid weakening signal-to-noise and therefore correlated sources could hypothetically still affect our results, but we did not see any evidence of correlated source artifacts. After notch filtering ~ 60 Hz, we filtered the data into four

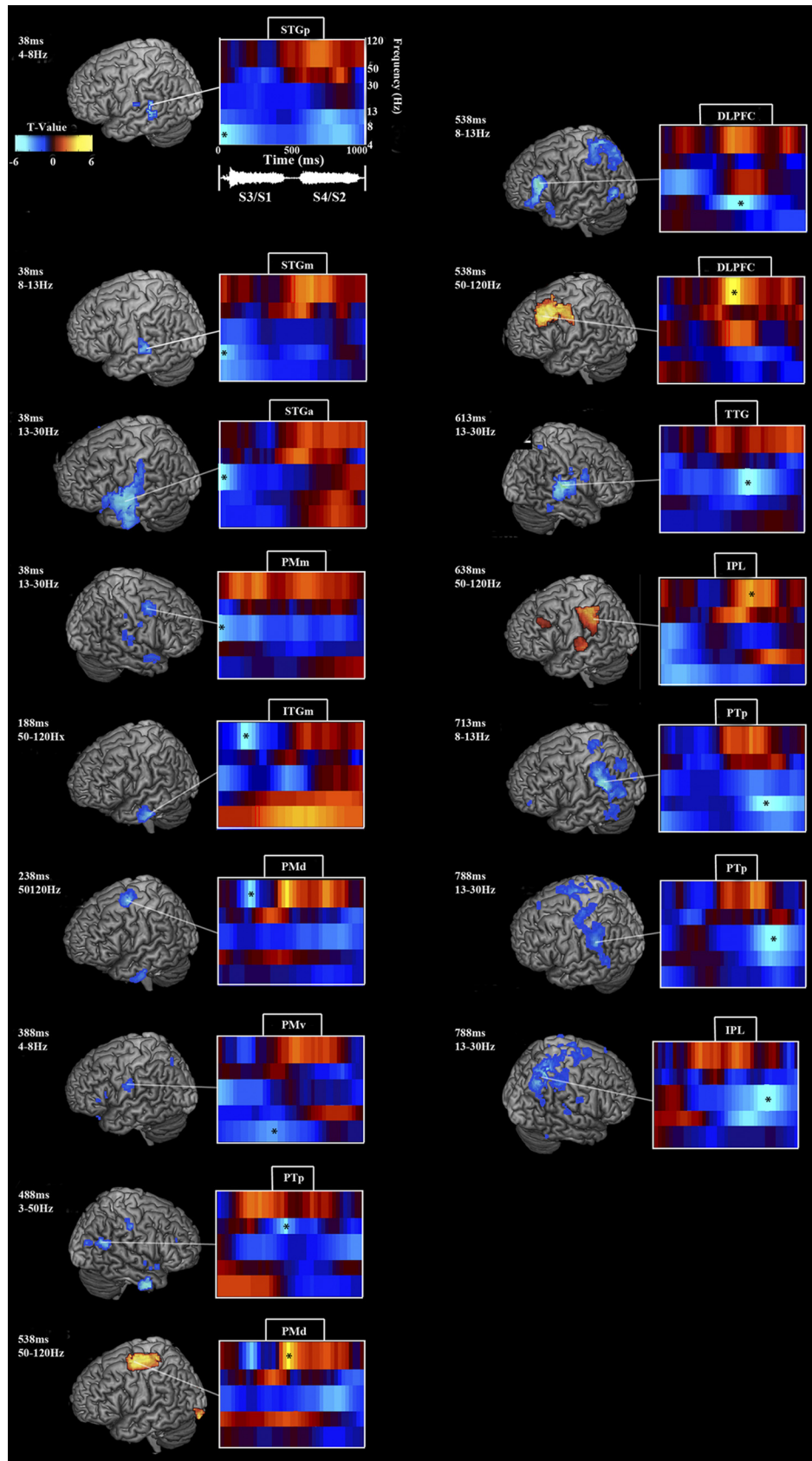


Figure 4. Effects of increased syllable load during encoding: oscillatory power contrast between second and first set of syllables on four-syllable trials. Figure layout follows Figure 2. Time 0 marks third/first syllable onset; fourth/second syllable onset occurs at ~520 ms. Low frequency (θ , α , and β) power was attenuated from the first to third syllable and the second to fourth syllable bilaterally over auditory and premotor areas. $H\gamma P$ increased over premotor and prefrontal areas from the second to the fourth syllable. For details, see Results and Table 3.

bands [4–13 Hz (θ/α), 13–30 Hz (β), 30–50 Hz (low gamma [$L\gamma$]), and 50–120 Hz (high gamma [$H\gamma$])] with a 60 and 120 Hz 1.5-Hz notch filter. For the load-effect analyses, we subdivided the lowest frequency band into θ (4–8 Hz) and α (8–13 Hz). Induced, phase-independent activity in each band was localized to subjects spatially normalized MRIs using the NUTMEG time–frequency beam-forming spatially adaptive filter algorithm, which has been described in detail previously (Dalal et al., 2011). In brief, for each 5 mm voxel, we computed a lead field describing the magnetic field strength at each sensor arising from a dipole source at the voxel. A time–frequency optimized beam-forming inverse solution for the dipole moment depending on the lead field and sensor covariance was then computed for each voxel for each frequency band at every time window, averaged across overlapping time windows. Localizations were computed using the shared computing cluster at the California Institute for Quantitative Biomedical Research (www.qb3.org). For activations, noise-corrected pseudo- F ratios were computed between active windows (i.e., subjects verbal response) and a prestimulus control baseline. For contrasts between the four- and two-syllable conditions, noise-corrected pseudo- F ratios were computed for sliding windows covering the time periods of interest. Window sizes were frequency-band optimized (4–13 Hz: 400 ms; 4–8 Hz: 400 ms; 8–13 Hz: 300 ms; 13–30 Hz: 200 ms; 30–50 Hz: 150 ms; 50–120 Hz: 100 ms) with an overlap of 25 ms. Activations were computed from averaged single-trial data covariance for each time window and frequency band. Separate analyses were performed for averages time locked to stimulus, visual go cue and response onset, to isolate onset latency and peak level of modulations in the activations. We analyzed only correct trials and balanced the number of two- and four-syllable trials for the four- versus two-syllable contrasts. For the stimulus, go cue, and response activation conditions, we grouped together two- and four-syllable trials to increase power.

Subjects were included/excluded from analysis strictly based on performance, but we applied different performance thresholds for inclusion in the power averages and the neurobehavioral regressions. The cutoff for subject inclusion in the power analysis was 90 total trials (50% accuracy), based on guidelines from previous work on tradeoffs between spatial resolution and signal detection in adaptive spatial filters (Sekihara et al., 2004; Brookes et al., 2008) and to ensure that the same group of subjects were maintained for the analysis of all phases of the experiment, thereby preventing any subject inclusion bias. For the correlation analysis, we wanted to use as broad a range of performance as possible and therefore chose to include more subjects in this analysis, also to ensure that our sensitivity to detect brain-behavior correlations was not underpowered as a result of sample size. For this analysis, we used a performance threshold for subject inclusion, namely 40% accuracy, which resulted in the exclusions of two subjects from the original 17 who completed the experiment. Although this threshold resulted in a lower amount of total data, the estimated effect on beam-forming accuracy and signal strength is of small enough magnitude so as to not change the regression result. Applying the same threshold for both the power analysis and the neurobehavioral regressions did not change our results qualitatively.

Group analyses were performed with statistical nonparametric mapping (SnPM) (Singh et al., 2003). The detailed rationale and procedures of SnPM statistics of beam-former images have been described previously (Singh et al., 2003; Dalal et al., 2008). In short, time–frequency beam-former images for each subject were first spatially normalized to the MNI template brain. The three-dimensional average and variance maps across subjects were calculated for each time–frequency window, and variance maps were smoothed with a $20 \times 20 \times 20$ mm³ Gaussian kernel. From this image, a pseudo- t statistic was obtained at each voxel, time window, and frequency band. Nonparametric null distributions were created by permuting voxel labels (2^N permutation, where N is the number of subjects) to derive p values for the true image that were then corrected for multiple comparisons across all voxels, frequency bands, and time points using the false discovery rate (FDR) procedure (Benjamini and Yekutieli, 2001). To assess neural–behavioral correlations, Pearson's correlation coefficients were computed for activations/contrasts for all voxels against reaction time and/or measures of accuracy. p values were corrected with the FDR procedure, using a threshold of 5%.

Statistical tests of performance were performed on rationalized arcsine transformed (normalized) data (Studebaker, 1985).

For the input and output buffer analyses, we regressed subjects' time–frequency activations and contrasts against behavior. We defined a metric that captures performance changes attributable to increased cognitive load, $SLP = (1 - (C_2 - C_4)/(C_2 + C_4))$, where C_2 is the number of correct two-syllable trials, and C_4 is the number of correct four-syllable trials. Because subjects always performed better on the two- than four-syllable trials, a high value for this syllable load performance metric (SLP) represents maintained performance from the two-syllable to four-syllable condition, normalized by overall performance. Thus, subjects that maintained performance on the more difficult four-syllable trials have similar SLP values, independent of their overall performance. For the input buffer analysis, we regressed activity during the first two syllables with total accuracy, because we grouped together correct and incorrect trials and contrast maps between the fourth and second syllable with four-syllable accuracy and with SLP (see Fig. 8A,B). Contrasting the fourth and second syllable as opposed to the first syllable allowed us to avoid confounding speech onset effects and provided a test more comparable with the cognitive load effect difference between the two- and four-syllable trials. To localize and examine output/production buffer effects, we regressed $H\gamma P$ in the response condition with accuracy (see Fig. 8C) and $H\gamma P$ contrast maps between the four- and two-syllable preproduction phases against SLP (see Fig. 8D).

Table 2. Activations during pre-response

Time (ms)	Band (Hz)	Area	MNI	MNI coordinates			t value	p value
				x	y	z		
–863	4–13	L PMv	BA6	–68	–5	18	–2.3	0.002
–813	50–120	L Pre-SMA	BA6	–3	18	68	2.7	0.001
–688	50–120	L PTr	BA45	–48	20	5	3.1	0.001
–638	13–30	L PMv	BA6	–63	3	35	–5.7	0.002
–538	50–120	L POs	BA9/44	–55	15	33	3.8	0.001
–463	13–30	R STSm	BA22	53	–48	0	–3.78	0.002
–463	13–30	L POs	BA44	–65	8	23	–5.2	0.002
–263	50–120	L PTr	BA45	–45	43	0	5.4	0.001
–213	50–120	R PMm	BA9	35	5	25	5.6	0.002
–213	50–120	L CSv	BA4	–53	–10	25	4.1	0.001
–188	13–30	L PMv	BA6	–55	–5	45	–5	0.002
–188	13–30	R CSv	BA43	63	–5	18	–3.1	0.002

Format follows Table 1.

Table 3. Encoding syllable load effects

Time (ms)	Band (Hz)	Area	MNI	MNI coordinates			t value	p value
				x	y	z		
38	4–8	L STGp	BA22	50	–33	10	–4.8	0.002
38	8–13	L STGm	BA22	–65	–30	8	–4.3	0.002
38	13–30	L STGa	BA22	–55	–3	0	–4.9	0.002
38	13–30	R PMm	BA6	35	0	48	–4.1	0.002
188	50–120	L ITGm	BA20	–63	–18	–28	–4.7	0.002
238	50–120	L PMd	BA6	–25	–5	63	–4.5	0.002
388	4–8	L PMv	BA6	–58	–10	30	–3.5	0.002
488	50–120	L PMd	BA6	–30	–3	63	4	0.001
538	8–13	L DLPFC	BA46	–45	35	20	–5	0.002
538	50–120	L DLPFC	BA9	–45	38	43	4.5	0.001
613	13–30	R TTG	BA41	45	–28	3	–5.1	0.002
638	50–120	L IPL	BA40	–59	–33	50	3.4	0.001
713	8–13	L PTP	BA40	–55	–45	23	–5.2	0.001
788	13–30	L PTP	BA40	–53	–40	38	–4.9	0.002
788	13–30	R IPL	BA7	30	–50	60	–5.5	0.002

Format follows Table 1.

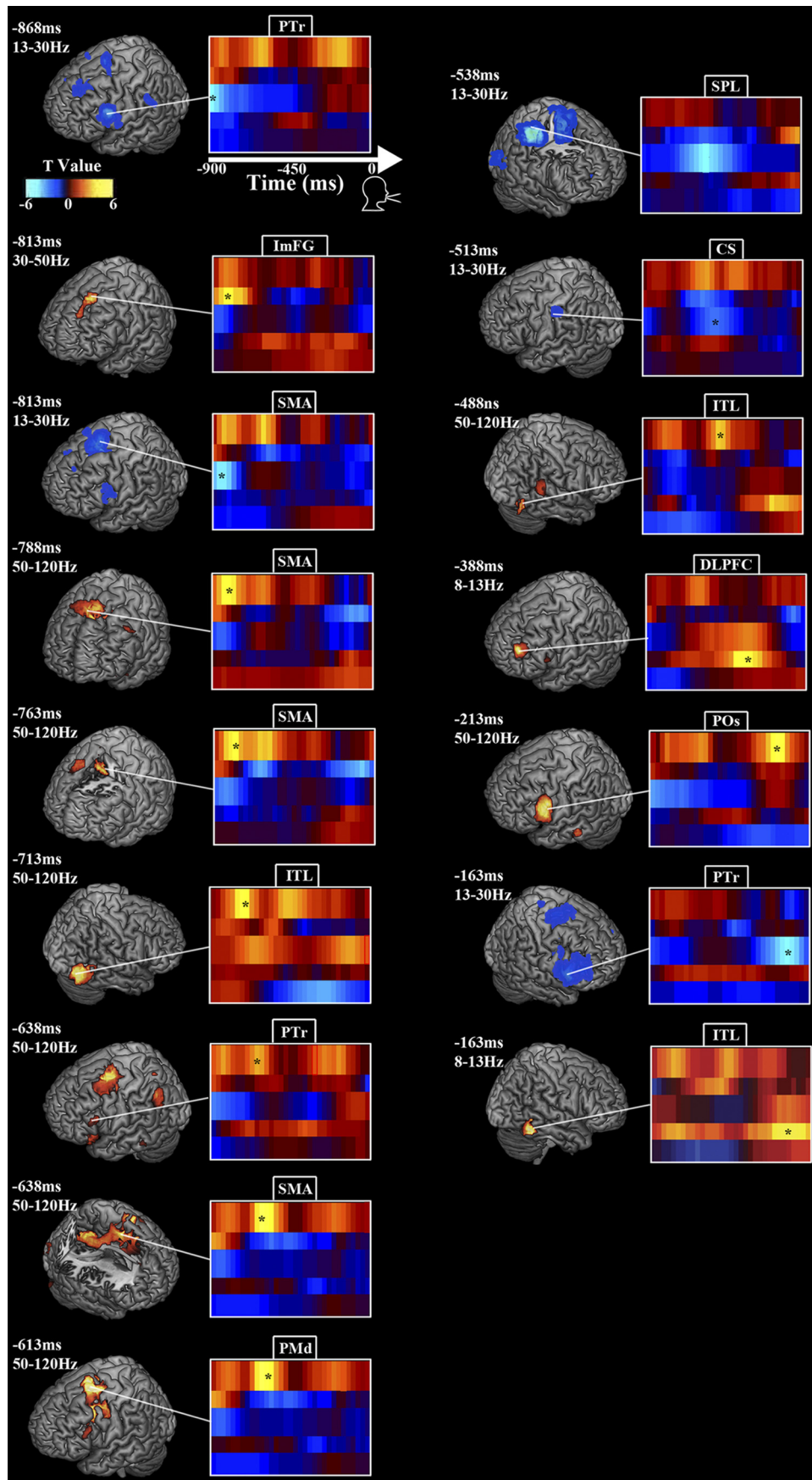


Figure 5. Effect of increased syllable load during speech preparation: oscillatory power contrast between four and two syllable trials. Figure layout follows Figure 2. Time 0 marks voice onset. β power contrast decreased and γ P contrast increased over SMA and PTR early in the pre-response period and POs/PMv late in the pre-response period. For details, see Results and Table 4.

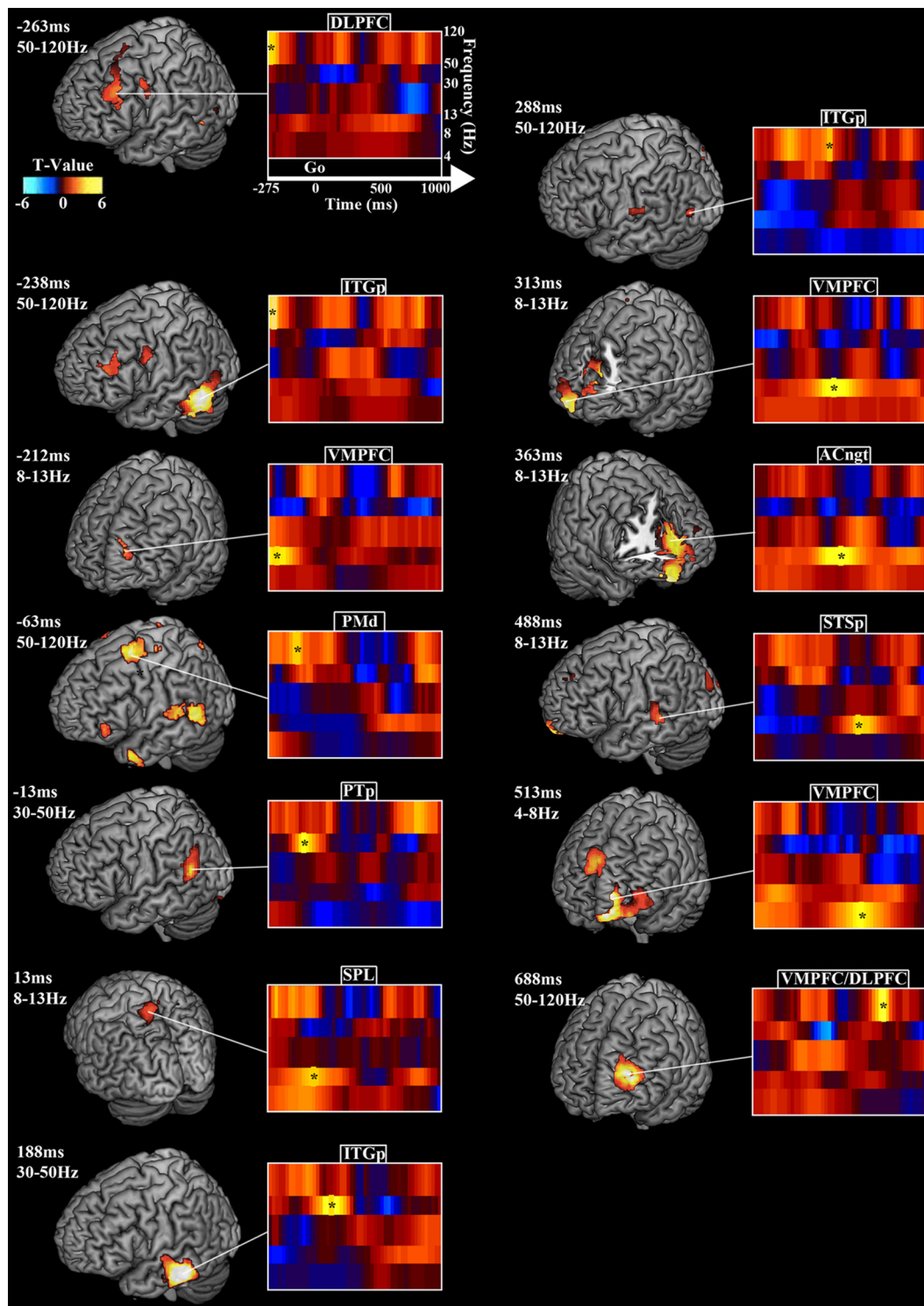


Figure 6. Effect of increased syllable load in the peri-go cue period: oscillatory power contrast between four and two syllable trials. Figure layout follows Figure 2. Time 0 marks the go cue. Only positive power contrasts from the two- to four-syllable condition emerged as significant. During the memory maintenance period preceding the go cue, $H\gamma P$ increased in prefrontal, precentral, and temporal areas. After the go cue, α and $H\gamma$ power increased over parietal, medial prefrontal, and posterior temporal areas. For details, see Results and Table 6.

Results

Behavioral responses

Subjects' accuracy and reaction time varied with task difficulty. Subjects performed significantly better on the two-syllable than the four-syllable trials ($p < 0.0001$) (Fig. 1B). Subjects repeated the two-syllable pattern correctly an average of 91% of the time

(73 ± 1.5 trials of 80) and the four-syllable pattern correctly an average of 63% of the time (50 ± 4.4 trials of 80). Reaction time was assessed as the average time between go cue and voice onset for each subject, using correct trials only. Reaction times for the two-syllable repetition (0.64 ± 0.03 s) were significantly lower than for the four-syllable repetition (0.83 ± 0.05 s) ($p < 0.001$) (Fig. 1C).

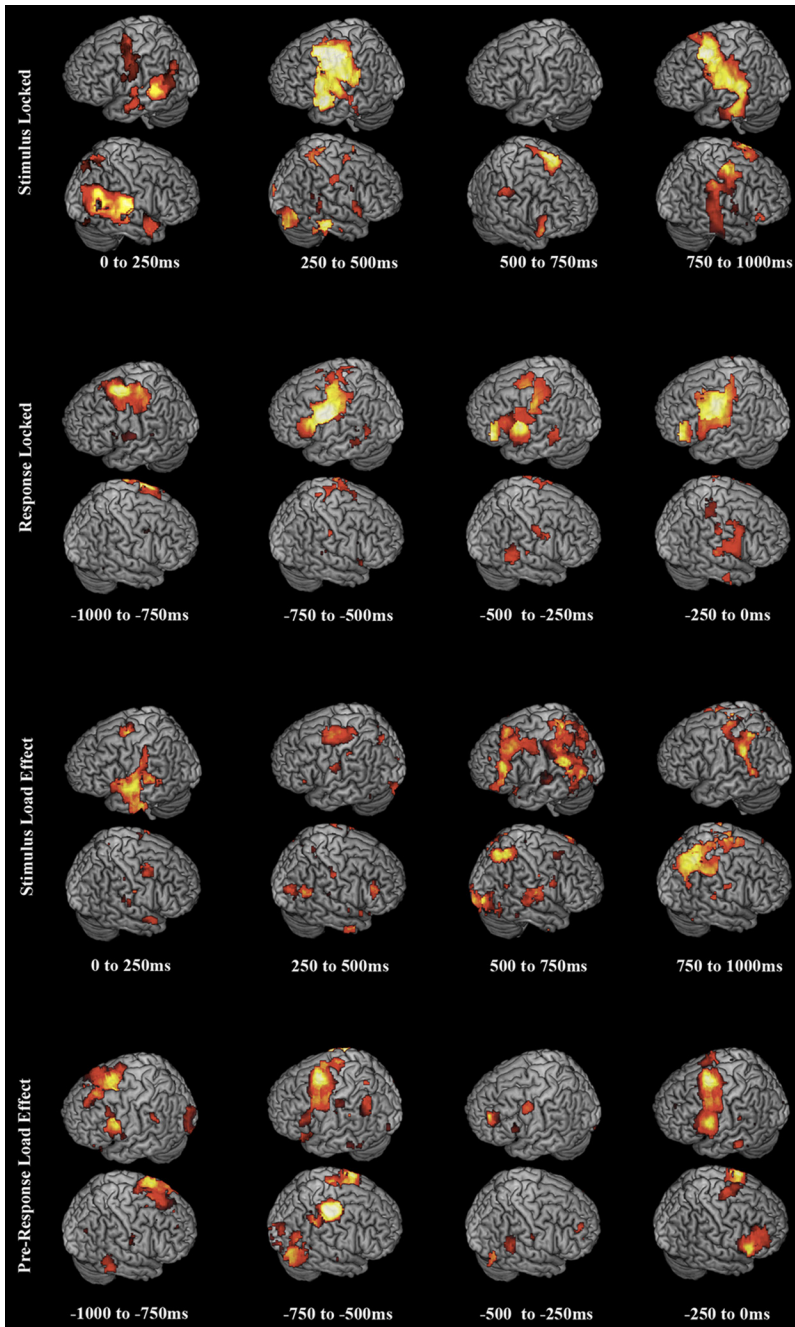


Figure 7. Summary figure depicting phonological loop network activations across all conditions. Brain renderings depict all θ , α , and β power decreases and $H\gamma P$ increases overlaid as activations, with absolute t values rendered additively. Activations and load effects oscillated between frontal and posterior regions of the dorsal speech stream. See Results and Discussion.

Neural activity

We first reconstructed oscillatory neural activity during the stimulus and pre-response periods, computing averages with respect to a prestimulus baseline for both two- and four-syllable trials together. Subsequently, we contrasted two- and four-syllable trials during the stimulus, peri-go cue, and pre-response periods, revealing differential activity across time and frequency associated with increased cognitive load. All peak time windows for statistically significant main effects and contrasts in our results are displayed in the figures. Finally, we describe oscillatory changes that are correlated with task performance across subjects, confirming that spectral power fluctuations represent be-

haviorally relevant cognitive processes. For all renderings, peak activations have corrected $p < 0.05$. Peak activations for significant activations (corrected $p < 0.05$) for all areas, time points, and conditions, along with their corresponding t and uncorrected p values, are listed in Tables 1–6.

Oscillatory power fluctuations during syllable encoding

Early auditory cortical responses to the syllable presentation manifested in temporally and spatially broad low-frequency (θ/α) power increases bilaterally, peaking at 37.5 ms post-sound onset in the left posterior STS (STSp) and 112.5 ms in the right supramarginal gyrus (SMG) (Fig. 2; Table 1). Accompanying the early auditory evoked field response, $H\gamma P$ increased over left PMv from 37.5 ms, and $L\gamma P$ decreased over left medial STG (STGm), peaking at 67.5 ms. After these early responses, α/θ power remained significantly elevated over left STG until 250 ms and over right STG until 375 ms and then peaked again over right posterior STG (STGp)/PTp at 837.5 ms during the second syllable. β power decreased bilaterally over temporal and frontal areas, peaking at 262.5 ms in left dorsal premotor cortex (PMd) and left anterior STG (STGa), right inferior temporal lobe (ITL) and inferior frontal gyrus (IFG) at 262.5 ms, and again in PMd at 862.5 ms and right medial central sulcus (CSm) at 837.5 ms. $L\gamma P$ decreased over left medial premotor cortex (PMm), peaking at 487.5 ms and again at 862.5 ms, and in right post-central gyrus (PsCG) at 200–1225 ms and PMd at 837.5 ms. $H\gamma P$ increased over right superior parietal lobe (SPL) at 162.5 ms, left CSm at 212.5 ms, right medial cingulate gyrus (MCngt) at 337.5 ms, right supplementary motor area (SMA) at 462.5 ms, and left PMm at 962.5 ms.

Oscillatory power fluctuations preceding speech production

Analysis of neural activity time locked to, but occurring before, voice onset (time 0) uncovered a network subserving reaction-time-independent processes (Fig. 3; Table 2), dominated by power changes in the β and $H\gamma$ bands. Early in the pre-response phase, θ/α power decreased over left PMv, peaking at 862.5 ms. In the left hemisphere, β power decreased over motor and premotor cortex, peaking over left PMv at -637.5 ms (relative to voice onset), pars opercularis (POs) at -462.5 ms, and again in PMm at -212.5 ms. $H\gamma P$ increased in pre-SMA/SMA at -812.5 ms, pars triangularis (PTr) at -687.5 ms, POs at -537.5 and -263 ms, and ventral central sulcus (CSv) at -212.5 ms. On the right, β power decreased in medial STS

(STSm) at -462.5 ms and CSv at -187.5 ms, and H γ P increased over PMm at -212.5 ms.

Oscillatory power fluctuations coding syllable memory load during stimulus encoding

To analyze the effect of syllable load during encoding, we contrasted oscillatory power during the second set of syllables with the first set of syllables (Fig. 4; Table 3). To fully explore the cognitive load effect, we subdivided the 4–13 Hz frequency band into θ (4–8 Hz) and α (8–13 Hz) separately. Times given are relative to the beginning of the third/first syllable contrast (0 ms) and extend out to the end of the fourth/second syllable contrast (1000 ms). Sound onset for the fourth/second syllables occurs at ~ 520 ms. In the left hemisphere, relative to the start of the third/first syllable contrast, θ power decreased over STGp, peaking at 37.5 ms, and PMv at 387.5 ms. α power reached a negative peak in STGm at 37.5 ms, dorsolateral prefrontal cortex (DLPFC) at 537.5 ms, and PTp at 712.5 ms. β power decreased over SMG/PTp at 787.5 ms. H γ P increased over medial inferior temporal gyrus (ITGm) at 187.5 ms, PMd at 237.5 ms, PMd at 487.5 ms, DLPFC at 537.5 ms, and IPL/PTp at 637.5 ms. This final H γ P load effect during the fourth syllable, arcing from IPL to STG, reflects an increase from and return to a relatively flat baseline throughout the first three syllables. In the right hemisphere, consistent effects were only seen in the β and L γ bands. β power decreased over PMm at 37.5 ms, over TTG (transverse temporal gyrus) at 612.5 ms, and SPL at 787.5 ms. L γ P decreased over PTp at 487.5 ms.

Oscillatory power fluctuations coding syllable memory load during response preparation

To analyze the response-phase syllable cognitive load effect, we contrasted the 900 ms period before vocal response between the four- and two-syllable trials, time locked to voice onset (Fig. 5; Table 4). We did not find any significant θ power differences, and α power differed significantly only at one time point, -387.5 ms (relative to voice onset) in left DLPFC, in which it was greater in the four-syllable condition, and L γ power differed only at -812.5 ms in intermediate frontal gyrus (ImFG). β and H γ power showed a variety of differences. In the left hemisphere, β power decreased over left PTr at -867.5 ms, SMA at -812.5 ms, and CSm at -512.5 ms; H γ P increased in PTr, insular cortex, and SMA at -637.5 ms, PMd at -612.5 ms, and POs at -212.5 ms. In the right hemisphere, β power decreased in IPL at -537.5 ms, H γ P increased in SMA at -787.5 ms, ITL at -712 and -487.5 ms, PTp at -487.5 ms, and the lingual gyrus (LNgG) at -512.5 ms.

Oscillatory power fluctuations coding syllable memory load time locked to go cue

To gain another perspective on the cognitive load effect from increased syllable memory, we computed four/two-syllable contrasts time locked to the go cue and examined oscillatory power changes before and after the go cue (Fig. 6; Table 5). We examined from 300 ms after the stimulus and before the go cue until 700 ms after the go cue. Although the α through L γ bands showed primarily decreased power relative to a prestimulus baseline, only relative power increases from the two- to four-syllable conditions emerged as statistically significant.

The α and H γ bands showed the most syllable load effects. In the left hemisphere, α power increased in ventromedial prefrontal

cortex (VMPFC) 212.5 ms before the go cue (-212.5 ms relative to the go cue), SPL at 12.5 ms, anterior cingulate gyrus (ACngt) and STSp at 487.5 ms. H γ P increased in DLPFC at -262.5 ms, posterior ITG (ITGp) at -237.5 ms, PMd at -62.5 ms, ITGp at 287.5 ms, and the VMPFC/DLPFC border at 687.5 ms. In the right hemisphere, α power increased in VMPFC at 312.5 ms and ACngt at 362.5 ms. θ power exhibited a positive power peak along the VMPFC/DLPFC border at 512.5 ms, and L γ power increased at 362.5 ms in the ACngt.

As an aid, we summarize the results from the preceding sections in Figure 7. We treat θ , α , and β power decreases and H γ power increases as activations on equal footing by summing the absolute t value of all statistically significant activations for each voxel across 250 ms time windows for each analysis condition. The result depicts the relative activity changes in each region on a coarse-grained timescale across all phases of the experiment.

Table 4. Pre-response syllable load effects

Time (ms)	Band (Hz)	Area	MNI	MNI coordinates			t value	p value
				x	y	z		
-868	13–30	L PTr	BA9/45	-53	13	23	-5.2	0.002
-813	30–50	L ImFG	BA8	-25	35	58	4.2	0.001
-813	13–30	SMA	BA6	-13	13	70	-4.8	0.001
-788	50–120	SMA	BA24	-8	0	43	4	0.001
-763	50–120	SMA	BA24	-3	0	50	4.2	0.001
-713	50–120	R ITL	BA19	53	-75	-5	4.6	0.001
-638	50–120	L PTr	BA45	-60	28	25	3.1	0.001
-638	50–120	SMA	BA6	-20	-8	53	4.6	0.001
-613	50–120	L PMd	BA6	-28	5	55	5.7	0.001
-538	13–30	R SPL	BA40	63	-43	50	-4.9	0.002
-513	13–30	L CS	BA1	-45	-23	38	-3.1	0.002
-488	50–120	R PTp	BA40	58	-45	13	3.5	0.001
-488	50–120	R ITL	BA19	48	-65	5	4	0.002
-388	8–13	L DLPFC	BA10	-35	5	23	4.6	0.001
-213	50–120	L POs	BA45	-60	20	23	4.9	0.001
-163	13–30	R PTr	BA45	60	23	5	-5	0.002
-163	8–13	R ITL	BA20	65	-47	-20	5	0.001

Format follows Table 1.

Table 5. Peri-go cue load effects

Time (ms)	Band (Hz)	Area	MNI	MNI coordinates			t value	p value
				x	y	z		
-263	50–120	R Insula	BA13	38	-10	8	4	0.002
-263	50–120	L DLPFC	BA9	-38	30	30	3.7	0.002
-238	50–120	L ITGp	BA37	-43	-45	-28	6.2	0.002
-212	8–13	L VMPFC	BA10	-15	65	30	3.8	0.002
-63	50–120	L PMd	BA6	-18	20	60	4.7	0.002
-63	50–120	L STGp	BA22	-60	-48	8	3.5	0.002
-13	30–50	L PTp/SMG	BA40	-58	-55	23	4.3	0.002
13	8–13	L SPL	BA7	-15	-60	68	3.3	0.002
13	50–120	R PreCUN	BA31	2.5	-65	28	4.5	0.002
188	30–50	L ITGp	BA21	-65	-23	-18	5	0.002
288	50–120	L ITGp	BA21	-58	-50	-8	3.6	0.002
288	50–120	PTm	BA43	-53	-18	18	3.2	0.002
313	8–13	L VMPFC	BA32	0	45	5	4.5	0.002
363	8–13	L ACngt	BA32	-5	43	15	4.5	0.002
488	8–13	L STSp	BA22	-63	-30	5	3.7	0.002
513	4–8	L VMPFC	BA10	-18	65	3	5.2	0.002
513	4–8	R SFG	BA9	3	55	35	4.2	0.002
688	50–120	L VMPFC	BA1	-38	60	23	5.4	0.002

Format follows Table 1.

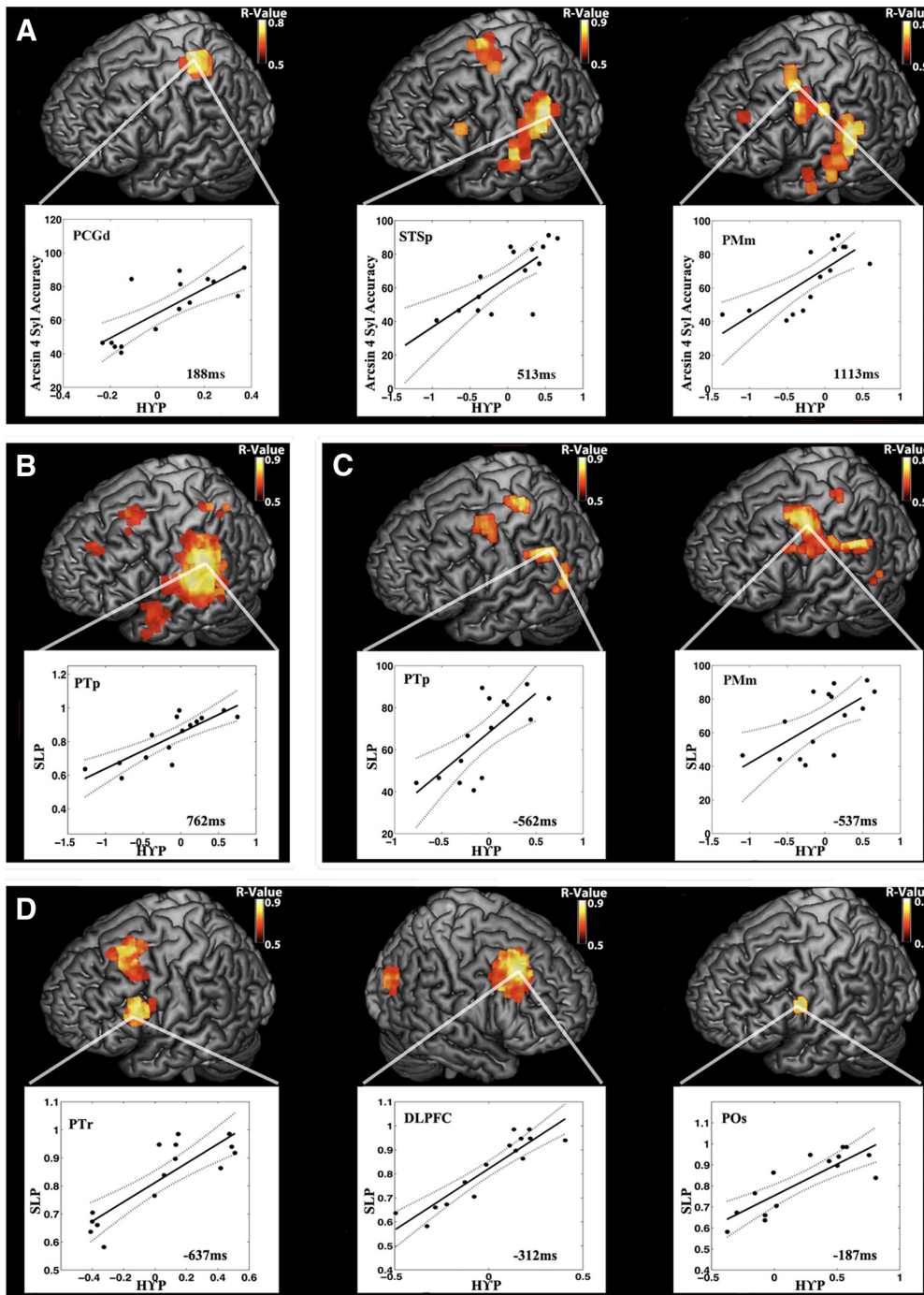


Figure 8. Peak H γ P performance correlations across stages of speech reproduction. For *A–D*, each brain rendering depicts the correlation at the time point indicated in the associated scatter plot, which depicts the individual subject data. In scatter plots, solid lines represent the best least-squares fit, and dotted lines represent the 95% confidence intervals. *A*, Stimulus period: peaks of correlation between whole-brain stimulus onset-locked H γ P and normalized four-syllable accuracy occur at 190 ms in PCGd, 525 ms in STSp, and 1125 ms after stimulus in PMm. *B*, Stimulus load effect: the whole-brain correlation between the four/two-syllable stimulus H γ contrast and the load performance metric peaks halfway through the syllable contrast at 262 ms after sound onset in PTP. *C*, Pre-response period: peaks of correlations between whole-brain response onset-locked H γ P and normalized total accuracy occur at 562 ms in PTP and 525 ms in PMm pre-response. *D*, Pre-response load effect: peaks of correlations between whole-brain four/two-syllable H γ contrast and load performance occur at 637 ms in left PTr, 312 ms in right DLPFC, and 187 ms in left POs pre-response.

Neural correlates of task performance

To further characterize the speech audiomotor network, we looked for correlations between subject performance and oscillatory power changes. For each condition (stimulus encoding, stimulus syllable load, pre-response, and response syllable load), we regressed the power in all voxels independently against subject accuracy, as well as the SLP that captures maintained perfor-

mance from two- to four-syllable conditions (see Materials and Methods). Neurobehavioral correlations emerged primarily within the dorsal speech stream, were mainly confined to the H γ band, and were primarily left lateralized.

During the encoding period of the four-syllable trials, in the left hemisphere, H γ P relative to the prestimulus baseline in STSp and posterior middle temporal gyrus (MTGp) exhibited a sus-

tained correlation with accuracy throughout the entire stimulus (Fig. 8A; Table 6). HγP in STSp also correlated with accuracy during most of the stimulus, peaking at 512.5 ms after stimulus onset and again at 1712.5 ms. The Hγ–accuracy correlation peaked in PMd at 187.5 ms after stimulus onset and again at 587.5 ms and in PMm at 1112.5 ms after stimulus, during the third

syllable. In the right hemisphere, the early θ power increase (corresponding to the auditory evoked field/m100) correlated with total accuracy. To explore the load effect in encoding, we contrasted the third and fourth syllable with the first and second, respectively, on the four phoneme trials and looked for correlation with SLP. A similar network of regions, including PTP and PMm, emerged as relevant for load-related performance, with the most significant correlations occurring for maintenance or increase of HγP from the second to the fourth syllable for both regions, at 762.5 ms relative to the second-syllable set versus first-syllable set contrast start point.

Table 6. HγP behavioral correlations

Analysis	Time	Band	Area	BA	MNI coordinates			R ²
					x	y	z	
Stimulus								
encoding								
	188	50–120	LPCGd	BA4	-32	-32	64	0.65
	513	50–120	LSTSp/SMG	BA40	-56	-40	24	0.64
	1113	50–120	LPm	BA6	-45	5	55	0.61
	1213	50–120	LSTSp	BA22	-56	-40	8	0.58
	1713	50–120	LPTp	BA22	-45	-32	10	0.7
Stimulus load								
	288	50–120	LMTGp	BA21	37	-55	-65	0.58
	338	50–120	RSPL	BA7	28	-65	53	0.79
	763	50–120	LPTp	BA22	-52	-38	17	0.57
Pre-response								
	-563	50–120	LPTp/SMG	BA40	-55	-45	30	0.62
	-538	50–120	LPm	BA6	-45	0	55	0.62
Response load								
	-637	50–120	LPTp	BA45	-60	20	25	0.58
	-560	50–120	PreSMA	BA6	-11	10	53	0.63
	-187	50–120	LPOs	BA44	-60	10	25	0.58
	-312	50–120	RDLPFC	BA8	40	30	50	0.84

Format follows Table 1.

During the pre-response period, HγP increases relative to the prestimulus baseline correlated early on with accuracy in PTP as well as in SMA (Fig. 8C), peaking at 562.5 ms before voice onset and in PMm at -537.5 ms. Closer to voice onset, positive HγP–accuracy correlations in POs trended toward but did not reach statistical significance after correcting for multiple comparisons. Examining the load effect during pre-production revealed that some of the areas exhibiting the most significant cognitive load effect also independently showed the most significant correlation with SLP (Fig. 8D). HγP increases from the two- to four-syllable trials correlated with SLP at left PTr at -637.5 ms and in left POs at -212.5 ms. HγP correlated highly with SLP in right DLPFC, peaking at -312.5 ms. Because the average HγP contrast peak in Broca’s area occurred 300 ms before speech onset but a difference wave extended until speech onset, this finding suggests that correct representation of a longer syllable pattern reflects in both persistent HγP and increased peak power. SLP also correlated with the shifted Hγ syllable load contrast time course in PMv and Broca’s area, confirming that earlier onset of Hγ modulation,

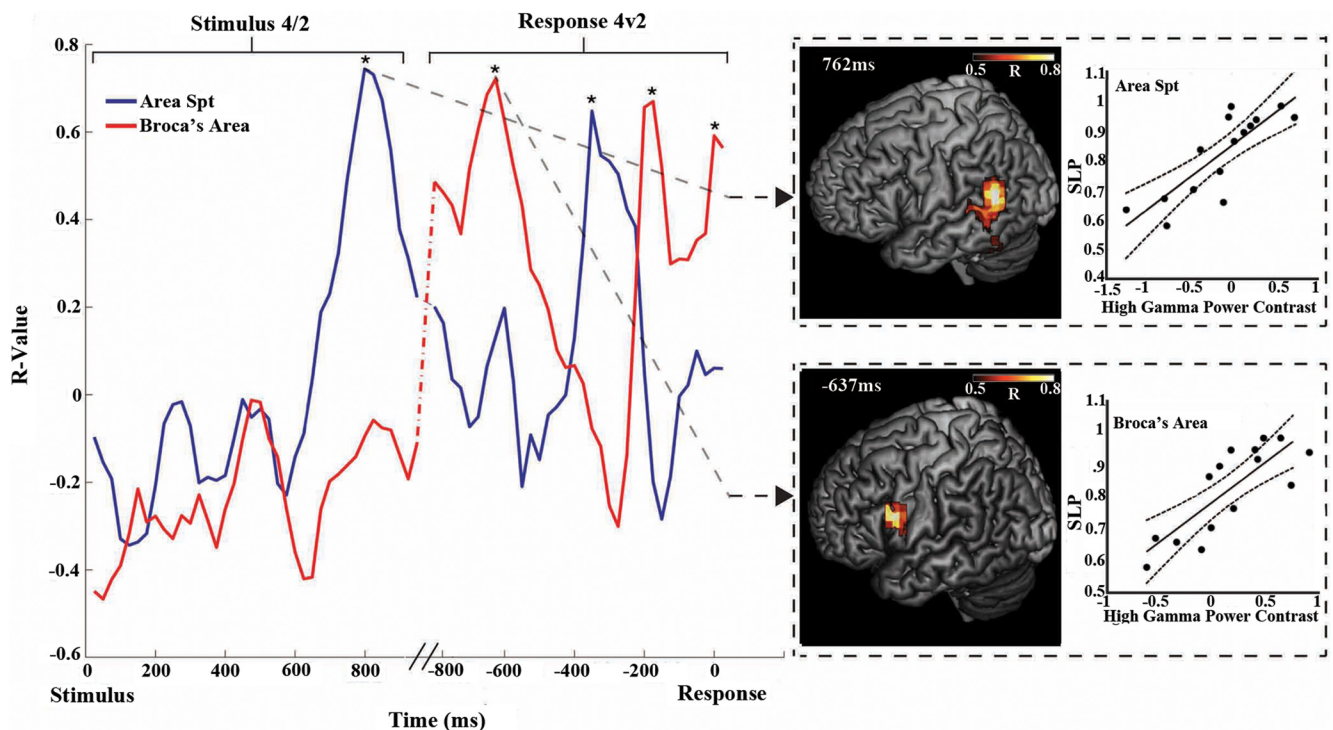


Figure 9. Reverberation in neurobehavioral correlation between area Spt and Broca’s area. Time courses of the correlations between HγP modulation and articulatory load-related performance (SLP) in area Spt/PTp (blue) and PMv/Broca’s area (red). Asterisks indicate statistically significant (FDR corrected, $*p < 0.05$) time points. Correlation time courses were computed separately for the stimulus and response periods and concatenated together for visualization. The stimulus contrast period begins at 0 ms, and the response period begins at -800 ms, with a 100 ms average time gap between the stimulus end and response period beginning (the variable gap represented by slanted lines). Whole-brain cortical localizations with accompanying scatter plots for Broca’s area and area Spt are shown for two time points, one during the stimulus period (the 4th syllable contrasted with the 2nd) and one during the response period (4-syllable trials contrasted with 2-syllable trials). Filled circles in scatter plots depict the individual data, solid lines the best linear fit, and dotted lines the 95% confidence intervals.

higher peak power, and a persistent upmodulation reflected in a smaller difference in the two- and four-syllable error rates.

Concatenating the stimulus and response-locked load effect correlation time courses reveals that articulatory-load-related activity alternates between Spt and Broca's area (Fig. 9). During the encoding period, the peak H γ P–SLP correlation localized to Spt, whereas in the pre-response phase the peak correlation localized to Broca's area. However, after the initial Broca's area *R* value peak, Spt again becomes important transiently for load-related performance at 300 ms before voice onset, only to trade off again with Broca's area at –200 ms.

Akin to the summary for the oscillatory power alone conditions depicted in Figure 7, in Figure 10, we summarize the results from all neurobehavioral correlation analyses.

Discussion

Our study revealed that speech reproduction success varies with the precise timing and frequency of oscillatory power shifts throughout the dorsal speech stream. We found that perceived speech target formation and motor plan development couple together in a reverberant process looping through frontal and posterior regions of the dorsal stream. While the lack of temporal resolution in previous studies has limited the dissociation of phonological loop functional components among brain regions, we show in detail how the pattern and behavioral relevance of cortical activations evolves throughout speech perception, working memory, and leading up to reproduction.

A left-lateralized subset of this overall speech–motor network responds to increased cognitive load, particularly in the H γ band. The activation levels of left STGp and IPL, including Spt, remained consistent during encoding of the first three syllables of the longer four-syllable stimulus, suggesting that the developing memory trace requires persistent but not increasing levels of activation. However, the marked increase in the left IPL and STGp H γ P from the second to the fourth syllable (Fig. 4, 638 ms), tapering off toward the end of the stimulus, suggests the performance of an additional computation, possibly an audiomotor transformation of the signal, as the entire sequence registers.

In addition to the more salient H γ effects, increased syllable-load induced changes across the entire frequency range. Consistent with the recent working memory literature, we found that increasing the number of syllables to be repeated resulted in a

complex picture of θ , α , β , and L γ band modulation. In seeming contrast with recent results on visual working memory (Palva et al., 2011), increased syllable load resulted in greater power decreases in putative task-relevant areas during ongoing encoding as well as response preparation. However, our experiment contained only a brief pure maintenance phase comparable with working memory studies, and during this period, only relative power increases emerged as significant, consistent with Palva et al. (2011). The pattern of relative low-frequency power decreases and increases during active encoding and response preparation, coupled with the more localized high-frequency power increases we found likely represents greater resource shifting from task-irrelevant to task-relevant areas in compensation for increasing cognitive load. This interpretation motivates Figure 7, in which H γ P increases and θ , α , and β decreases are overlaid with equal weight as activations. Our results suggest further, as seen in STG/IPL in our study, that areas may be alternatively activated and deactivated on a timescale too fast for evoked and temporally coarse induced reconstructions. We saw trending evidence in auditory and frontal/motor areas of alternation in relative power between low- and high-frequency bands, suggesting that both activating and deactivating processes play roles in speech audiomotor transformation.

We found two H γ load-effect peaks in the pre-response phase, indicating the possibility of two levels of representation of the speech target. The earlier more rostral peak in PTr and left insular cortex, synchronous with SMA activation, corresponds to the longer reaction time for four-syllable trials and could reflect the more complicated action selection required for framing and ordering the four-syllable utterance; the later more caudal load effect in POs likely results from the simultaneous representation of syllables in an output buffer, consistent with competitive queuing models, such as GODIVA (Bohland et al., 2010). This gradient reflects the overall rostral–caudal hierarchical organization of frontal areas and Broca's area in particular (Gelfand and Bookheimer, 2003; Badre and D'Esposito, 2009; Sahin et al., 2009) and suggests that auditory areas may communicate different features of speech sequences to rostral Broca's area and caudal Broca's area/PMv, respectively, consistent with evidence of differential white matter connectivity and possibly reflecting the division between the ventral and dorsal speech streams (Brauer et al., 2011; Ueno et al., 2011). However, greater spatial resolu-

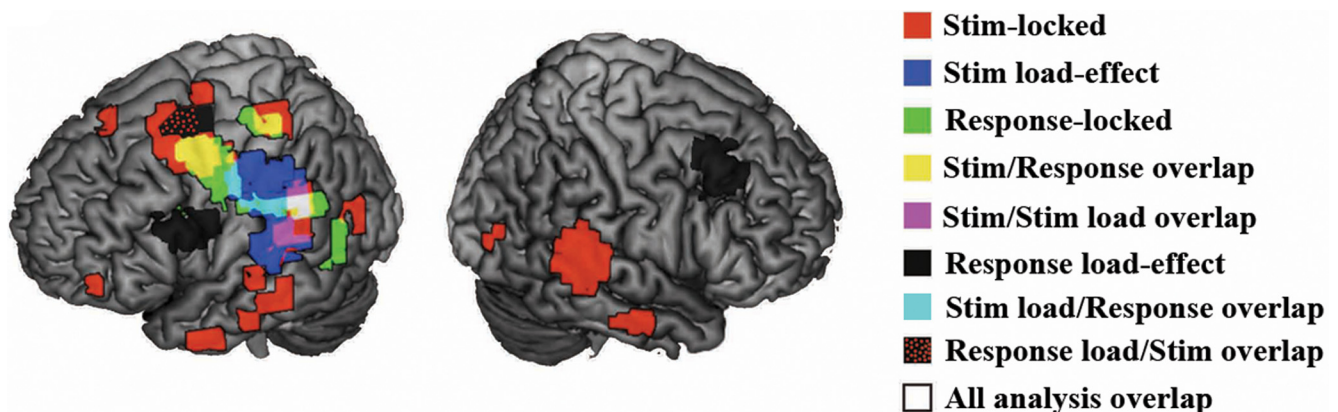


Figure 10. Combined overlays of peak H γ performance correlations for all analyses. Overlays thresholded above $R^2 = 0.6$. Red, Stimulus locked correlations; blue, stimulus load effect correlations; green, response locked correlations; dark gray, response load effect correlations; yellow, overlap between stimulus and response locked; magenta, overlap between stimulus-locked and stimulus load effect; cyan, overlap between stimulus load effect and response locked; black with red dots, overlap between stimulus-locked and response load effect; white, overlap between all conditions.

tion is required to definitively distinguish these subdivisions of Broca's area.

The time courses of both activations and correlations between oscillatory power and performance suggest that the dorsal stream phonological loop network operates dynamically during both encoding and pre-production. Information about the speech target to be repeated cycles through the dorsal stream, and representations are formed as part of an interactive process between frontal/premotor and temporal/inferior-parietal areas. During the encoding phase of our experiment, medial and superior temporal, PTP/IPL activation, motor, dorsal premotor, and inferior frontal areas all became engaged; however, the posterior areas dominated the behavioral correlations. The frontal and central sulcal activations possibly reflect the priming of the speech motor system and the initialization of a motor representation and also correlate significantly with performance (Yuen et al., 2010). Based on the activation timing and behavioral relevance, the PTP localization likely represents an average of the individual Spt locations, which are known to vary (Buchsbaum et al., 2011). The MTG and STS correlations likely correspond to phoneme mapping, whereas the PTP/IPL correlations correspond to phoneme storage and audiomotor transformation (Hickok et al., 2011). However, the spatial resolution of the present work does not allow us to rule out the possibility that audiomotor transformation and perceptual buffering occur at distinct loci within the STGp/IPL region. During the pre-production phase, Broca's area/PMv and pre-SMA trade off with Spt initially; however, within 250 ms of voice onset, frontal regions dominate the performance correlations, suggesting that by that time a motor representation of the entire syllable sequence has formed.

The dissociations that we observe between the neurobehavioral correlation time courses in Spt and Broca's area coupled with the time courses of activations add significant support to the notion that these areas act as buffers for storing syllable representations within the phonological loop and furthermore demonstrate how the developing speech target representations reverberate between these buffers. This cycling through the phonological loop suggests a more dynamic efference copy registration process than commonly believed, with the Broca's area buffer updating its representations in real time as it synchronizes with the short-term verbal memory representation in Spt, extending current state-space models of speech motor control (Houde and Nagarajan, 2011).

This bidirectional nature of the interaction between Spt and Broca's area implies that, in addition to fulfilling their functions as input and output buffers, at least one cycle of an internal feedback loop between the two must complete after stimulus encoding for successful production to occur. Such a feedback loop, possibly involving an efference copy of the upcoming speech output as suggested above, would represent a generalized version of Baddeley's articulatory rehearsal process (Baddeley, 2000). If the model our results suggest proves correct, then interfering with Spt at the end of an auditory stimulus should disproportionately affect the reproduction of longer utterances, whereas interfering after the go cue but before voice onset would affect short and long sequences equally. Similarly, targeting STS during early stimulus encoding should affect performance more than Spt, whereas deactivating Spt toward the end of encoding should disrupt repetition more than deactivating STS. Frontally, inhibiting Broca's area before speech onset would affect longer utterances before shorter ones, whereas deactivating Broca's area during the stimulus presentation should minimally affect performance of either

long or short sequences, consistent with the symptoms of expressive aphasias.

Collapsing across all frequency bands and task phases, the neuroanatomical network we found reproduces the function sparing in the rare clinical syndrome known as mixed transcortical aphasia (MTA) (Grossi et al., 1991; Yankovsky and Treves, 2002) and unifies and extends past functional imaging results (Dronkers, 1996; Hillis et al., 2005; Bohland and Guenther, 2006; Buchsbaum and D'Esposito, 2008; Towle et al., 2008; Papoutsis et al., 2009; Sahin et al., 2009; Edwards et al., 2010) (Figs. 2–4, 8). The speech abilities preserved in MTA map almost one-to-one onto the functions of the phonological loop and correspond closely with our putative phonological loop sites, including IPL, STG, PTP, PMd and PsCG, PMv, and Broca's area.

Furthermore, our results suggest that disruption in the timing of activation of Spt and its interaction with Broca's area and pre-SMA in this automatic pre-production feedback loop may contribute to the phonemic paraphasias observed in conduction aphasia and cognitive disorders, such as schizophrenia, Alzheimer's disease, and primary progressive aphasia (Wherrett, 2008). For instance, accumulating evidence points to Spt damage as the sine-qua-non of conduction aphasia (Buchsbaum et al., 2011). If this feedback loop after target perception/identification but before speech onset becomes disrupted, or the representation in Spt decays before the corresponding one in Broca's area solidifies, then spoken outputs would likely become "noisy," leading to the kind of error-prone paraphasic speech, coupled with the inability to correct speech errors, observed in conduction aphasia and some cases of schizophrenia (Gerson et al., 1977; Heinks-Maldonado et al., 2007; Ardila, 2010; Hickok et al., 2011; Lane et al., 2011).

In conclusion, we present a high-resolution spatiotemporal description of the brain network that supports the functions of the phonological loop; we demonstrate that Spt and Broca's area act as dissociable input and output buffers in a reverberating bidirectional dorsal stream for phoneme perception and production. These buffers interact in time along a tightly controlled schedule as part of a speech-motor feedback loop that operates before speech onset. Our results further indicate that treatments for clinical conditions that include symptoms of paraphasia should include a focus on the connectivity and timing of interaction between frontal and posterior areas, to facilitate the critical reinforcement of correct motor representations that occurs through input and output buffer internal model feedback.

References

- Ardila A (2010) A review of conduction aphasia. *Curr Neurol Neurosci Rep* 10:499–503. [CrossRef Medline](#)
- Baddeley A (2000) The episodic buffer: a new component of working memory? *Trends Cogn Sci* 4:417–423. [CrossRef Medline](#)
- Baddeley AD (2010) Working memory. *Curr Biol* 20:5.
- Baldo JV, Dronkers NF (2006) The role of inferior parietal and inferior frontal cortex in working memory. *Neuropsychology* 20:529–538. [CrossRef Medline](#)
- Barch DM (2005) The cognitive neuroscience of schizophrenia. *Annu Rev Clin Psychol* 1:321–353. [CrossRef Medline](#)
- Benjamini Y, Yekutieli Y (2001) The control of the false discovery rate under dependency. *Ann Stat* 29:1165–1188. [CrossRef](#)
- Bohland JW, Guenther FH (2006) An fMRI investigation of syllable sequence production. *Neuroimage* 32:821–841. [CrossRef Medline](#)
- Bohland JW, Bullock D, Guenther FH (2010) Neural representations and mechanisms for the performance of simple speech sequences. *J Cogn Neurosci* 22:1504–1529. [CrossRef Medline](#)

- Brauer J, Anwander A, Friederici AD (2011) Neuroanatomical prerequisites for language functions in the maturing brain. *Cereb Cortex* 21:459–466. [CrossRef Medline](#)
- Brookes MJ, Vrba J, Robinson SE, Stevenson CM, Peters AM, Barnes GR, Hillebrand A, Morris PG (2008) Optimising experimental design for MEG beamformer imaging. *Neuroimage* 39:1788–1802. [CrossRef Medline](#)
- Buchsbaum BR, D'Esposito M (2008) The search for the phonological store: from loop to convolution. *J Cogn Neurosci* 20:762–778. [CrossRef Medline](#)
- Buchsbaum BR, Baldo J, Okada K, Berman KF, Dronkers N, D'Esposito M, Hickok G (2011) Conduction aphasia, sensory-motor integration, and phonological short-term memory: an aggregate analysis of lesion and fMRI data. *Brain Lang* 119:119–128. [CrossRef Medline](#)
- Dalal SS, Guggisberg AG, Edwards E, Sekihara K, Findlay AM, Canolty RT, Berger MS, Knight RT, Barbaro NM, Kirsch HE, Nagarajan SS (2008) Five-dimensional neuroimaging: localization of the time-frequency dynamics of cortical activity. *Neuroimage* 40:1686–1700. [CrossRef Medline](#)
- Dalal SS, Zumer JM, Guggisberg AG, Trumpis M, Wong DDE, Sekihara K, Nagarajan SS (2011) MEG/EEG source reconstruction, statistical evaluation, and visualization with NUTMEG. *Comput Intell Neurosci* 2011:17. [CrossRef Medline](#)
- Badre D, D'Esposito M (2009) Is the rostro-caudal axis of the frontal lobe hierarchical? *Nat Rev Neurosci* 10:659–669. [CrossRef Medline](#)
- Dronkers NF (1996) A new brain region for coordinating speech articulation. *Nature* 384:159–161. [CrossRef Medline](#)
- Edwards E, Nagarajan SS, Dalal SS, Canolty RT, Kirsch HE, Barbaro NM, Knight RT (2010) Spatiotemporal imaging of cortical activation during verb generation and picture naming. *Neuroimage* 50:291–301. [CrossRef Medline](#)
- Fukuda M, Rothermel R, Juhász C, Nishida M, Sood S, Asano E (2010) Cortical gamma-oscillations modulated by listening and overt repetition of phonemes. *Neuroimage* 49:2735–2745. [CrossRef Medline](#)
- Gelfand JR, Bookheimer SY (2003) Dissociating neural mechanisms of temporal sequencing and processing phonemes gyrus. *Neuron* 38:831–842. [CrossRef Medline](#)
- Gerson SN, Benson F, Frazier SH (1977) Diagnosis: schizophrenia versus posterior aphasia. *Am J Psychiatry* 134:966–999. [Medline](#)
- Grossi D, Trojano L, Chiacchio L, Soricelli A, Mansi L, Postiglione A, Salvatore M (1991) Mixed transcortical aphasia: clinical features and neuroanatomical correlates. A possible role of the right hemisphere. *Eur Neurol* 31:204–211. [CrossRef Medline](#)
- Heinks-Maldonado TH, Mathalon DH, Houde JF, Gray M, Faustman WO, Ford JM (2007) Relationship of imprecise corollary discharge in schizophrenia to auditory hallucinations. *Arch Gen Psychiatry* 64:286–296. [CrossRef Medline](#)
- Hickok G, Poeppel D (2007) The cortical organization of speech processing. *Nat Rev Neurosci* 8:393–402. [CrossRef Medline](#)
- Hickok G, Houde J, Rong F (2011) Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* 69:407–422. [CrossRef Medline](#)
- Hillis AE, Work M, Barker PB, Jacobs MA, Breese EL (2005) Re-examining the brain regions crucial for orchestrating speech articulation. *Brain* 137:1479–1487. [CrossRef Medline](#)
- Houde JF, Nagarajan SS (2011) Speech production as state feedback control. *Front Hum Neurosci* 5:82. [CrossRef Medline](#)
- Jacquemot C, Dupoux E, Bachoud-Lévi AC (2007) Breaking the mirror: asymmetrical disconnection between the phonological input and output codes. *Cogn Neuropsychol* 24:3–22. [CrossRef Medline](#)
- Lane ZP, Singer A, Roffwarg DE, Messias E (2011) Differentiating psychosis versus fluent aphasia. *Clin Schizophr Relat Psychoses* 4:258–261. [CrossRef Medline](#)
- McGettigan C, Warren JE, Eisner F, Marshall CR, Shanmugalingam P, Scott SK (2011) Neural correlates of sublexical processing in phonological working memory. *J Cogn Neurosci* 23:961–977. [CrossRef Medline](#)
- Monsell S (1987) On the relation between lexical input and output pathways of speech. In: *Language perception and production: relationships between listening, speaking, reading and writing* (Allport A, MacKay DG, Prinz W, Scheerer E, eds), pp 273–311. London: Academic.
- Palva S, Kulashekhar S, Hämäläinen M, Palva JM (2011) Localization of cortical phase and amplitude dynamics during visual working memory encoding and retention. *J Neurosci* 31:5013–5025. [CrossRef Medline](#)
- Papoutsis M, de Zwart JA, Jansma JM, Pickering MJ, Bednar JA, Horwitz B (2009) From phonemes to articulatory codes: an fMRI study of Broca's area in speech production. *Cereb Cortex* 19:2156–2165. [CrossRef Medline](#)
- Rauschecker JP, Scott SK (2009) Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci* 12:718–724. [CrossRef Medline](#)
- Sahin NT, Pinker S, Cash SS, Schomer D, Halgren E (2009) Sequential processing of lexical, grammatical and phonological information within Broca's area. *Science* 326:445–449. [CrossRef Medline](#)
- Saur D, Kreher BW, Schnell S, Kümmerer D, Kellmeyer P, Vry MS, Umarova R, Musso M, Glauche V, Abel S, Huber W, Rijntjes M, Hennig J, Weiller C (2008) Ventral and dorsal pathways for language. *Proc Natl Acad Sci U S A* 105:18035–18040. [CrossRef Medline](#)
- Sekihara K, Nagarajan SS, Poeppel D, Marantz A (2004) Asymptotic SNR of scalar and vector minimum-variance beamformers for neuro-magnetic source reconstruction. *IEEE Trans Biomed Eng* 51:1726–1734. [CrossRef Medline](#)
- Singh KD, Barnes GR, Hillebrand A (2003) Group imaging of task-related changes in cortical synchronisation using nonparametric permutation testing. *Neuroimage* 19:1589–1601. [CrossRef Medline](#)
- Studebaker GA (1985) A “rationalized” arcsine transform. *J Speech Hear Res* 28:455–462. [Medline](#)
- Towle VL, Yoon HA, Castelle M, Edgar JC, Biassou NM, Frim DM, Spire JP, Kohrman MH (2008) ECoG gamma activity during a language task: differentiating expressive and receptive speech areas. *Brain* 131:2013–2027. [CrossRef Medline](#)
- Ueno T, Saito S, Rogers TT, Lambon Ralph MA (2011) Lichtheim 2: synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. *Neuron* 72:385–396. [CrossRef Medline](#)
- Vallar G (2006) The case of phonological short-term memory: A festschrift for cognitive neuropsychology. *Cogn Neuropsychology* 139:413.
- Wherrett JR (2008) The role of the neurologic examination in the diagnosis and categorization of dementia. *Geriatr Aging* 11:203–208.
- Yankovsky AE, Treves TA (2002) Postictal mixed transcortical aphasia. *Seizure* 11:278–279. [CrossRef Medline](#)
- Yuen I, Davis MH, Brysbaert M, Rastle K (2010) Activation of articulatory information in speech perception. *Proc Natl Acad Sci U S A* 107:592–597. [CrossRef Medline](#)